**Abstract.** *Market research through surveys uses several data collection methods, usually: face to face interviews (PAPI), telephone interviews (CATI), online interviews (CAWI). This paper presents the data collection methods currently used in the world and in Romania and a comparison between methods depending on the main criteria of choice: sample size, interview length, target incidence and survey price. Statistical analyses based on these variables lead to an evaluation of the methods on each criterion; this reveals situations where certain methods are more suitable than others, depending on which criterion is most important for the research. The analysis is based on actual price simulations in a marketing research company and is relevant for the current status of market research in Romania.*

**Keywords:** computer aided telephone interviews (CATI), computer aided web interviews (CAWI), data collection methods, market research through surveys, pen and paper interviews (PAPI).

# DATA COLLECTION IN ROMANIAN MARKET RESEARCH: A COMPARISON BETWEEN PRICES OF PAPI, CATI AND CAWI

**Alina BARBU**
*Academy of Economic Studies, Bucharest*
*Piața Romană no. 6, Bucharest, Romania*
*e-mail: av_barbu@yahoo.com*

**Alexandru ISAIC-MANIU**
*Academy of Economic Studies, Bucharest*
*Piața Romană no. 6, Bucharest, Romania*
*e-mail: al.isaic-maniu@csie.ase.ro*

*Management & Marketing*
*Challenges for the Knowledge Society*
*(2011) Vol. 6, No. 3, pp. 349-364*

## 1. Introduction

Quantitative market research uses surveys to analyze samples and extend results to the entire population. Several data collection methods are available, most commonly: face to face, telephone, online, and mail interviews. These are employed today throughout the world in various degrees, reflecting different stages of the evolution of society. *Face to face, paper interviews* are the chronological basis of survey research. The increased presence of telephones in households allowed the development of large-scale *telephone interviews* and technology enabled telephone interviews to be conducted on computers. The extended Internet usage nowadays makes *online surveys* possible, which, depending on Internet penetration, can be representative of certain targets.

**Market research in the world – what methods are used and where does it stand?**

The global development of survey methods heads towards *faster* and *less costly* data collection. These are the two requirements posed upon market research from survey beneficiaries, pressure which helped data collection methods move towards those used today. The data collection method used most often in a country is usually a good indicator of how advanced market research is in that country. Advanced methods are available, but their usage depends greatly upon the technology available in each country.

The Esomar report "Global Market Research 2009" shows that, in the year 2008, market research had a worldwide turnover of 32 billion dollars and that 80% of the research was conducted through quantitative methods. Europe had the lowest growth rate compared to the previous year: only 4.7% growth (adjusted with inflation). In contrast, emergent markets like Latin America, the Middle East and Africa had growths of 12-13%. Romania had an important growth as well – 22.8%, which shows that market research is not yet a mature field here either.

It was suggested that 50% of the worldwide market research budget would be spent on online research by 2005 (James, 2000). Only half of this estimate came true: *the most used data collection method in 2008 was online research, but it accounted for 20% of the total budget* (25% of quantitative research). It was followed by telephone research (18% of total and 23% of quantitative) and face to face research (12% of total and 15% of quantitative).

**Market research in Romania – what methods are used and where do we stand compared to other countries?**

The situation in Romania is slightly different. Market research has not reached maturity here, therefore *the most used data collection methods are face to face and telephone interviews* – PAPI (pen and paper interviews) and CATI (computer aided

telephone interviews). Online surveys – *computer aided web interviews (CAWI) account for only 1% of the total research*.

The difference is explained by technological development: in June 2009, the Internet penetration rate in households was 32.1% (ANCOM). This means that only one in three Romanian households are connected to the Internet, the situation being only little better in the urban environment (1 in 2 households are online). Under such circumstances, nationally representative online surveys cannot be conducted.

At the moment, technological development holds Romania back from full-scale CATI and CAWI research. While Romania is struggling with implementing online surveys, GfK NOP, U.K. market research institute, takes research one step further: they are running international surveys on several types of mobile phones (Strong, 2010). In conclusion, looking at market research in Romania compared to the world and advanced countries, we can notice *important development opportunities, both in terms of research spendings and in terms of methods used.*

## 2. Criteria for choosing a data collection method

Choosing a data collection method depends on several criteria. Tull and Hawkins (1990) propose 7 important factors in choosing the methodology, which refer to the statistical aspects of a survey, summarized in Table 1. Their outlook is similar to that of a research agency, aiming at delivering statistically significant results.

*Table 1*

**Evaluating data collection methods, after Tull and Hawkins**

| Criterion | CATI | PAPI | CAWI |
|---|---|---|---|
| 1. Handling complex questionnaires | Good | Excellent | Good |
| 2. Collecting large amounts of data | Good | Excellent | Good |
| 3a. Response accuracy in sensitive questions | Average | Average | Good |
| 3b. Control over interviewer effects | Average | Low | Excellent |
| 4. Sample control | Excellent | Average | Average |
| 5. Time needed for the research | Excellent | Good | Good |
| 6. Probable answer rate | Average | Average | Average |
| 7. Cost | Good | Average | Average |

On the other hand, Deblès and Teyssonnière de Gramont (1991) consider 8 factors important in choosing the data collection method, summarized in Table 2. These factors are similar to the previous ones, but they are more closely related to project management than to statistics. This is closer to the client's outlook on the research.

There is no universally valid answer for choosing the data collection method. The previous authors offered several possible options; common criteria include: *interview length, response rate and questionnaire flexibility/ complexity*.

As both authors note, *price* is an important factor whether the main criteria are statistical or managerial. Market research is a competitive field, where pressure from the beneficiary reflects in price reductions from the agency. Each method has advantages and disadvantages which ultimately affect the cost of the study and therefore are taken into account when choosing the methodology.

*Table 2*

**Evaluating data collection methods, after Deblès and Teyssonnière de Gramont**

| Criterion | CATI | PAPI | CAWI* |
|---|---|---|---|
| 1. Unitary cost | High | Average | High |
| 2. Time needed for the research | Average | Low | Low |
| 3. Control over respondents' reaction | Excellent | Good | Good |
| 4. Control over data collection process | Low | Average | Excellent |
| 5. Questionnaire flexibility | Excellent | Average | Average |
| 6. Geographic spread of respondents | Average | Good | Good |
| 7. Refusal rate and/or missing answers | Possibly high | Average | Average |
| 8. Respondents' perception on anonymity of answers | Possible problem | Average | Average |

** Authors included here video questionnaires, which are more complex visually and in terms of programming and therefore more expensive.*

In the following chapters we present a statistical analysis of costs of main data collection methods in Romania. There are mixed opinions regarding the quality of data obtained through various methods. Some research shows that there is no significant difference between the quality of PAPI and CATI (see endnote [1]), while others argue for the better quality of data coming from long PAPI interviews compared to CATI (see endnote [2]). The analysis takes into account only price and its factors of influence and assumes the quality of data to be similar throughout the cases analysed.

## 3. Factors which influence the price of PAPI, CATI and CAWI in Romania

Price is important in choosing the research methodology; therefore we analyzed the price of main methodologies used in Romania: PAPI, CATI and CAWI to see which factors influence it. Understanding what triggers significant price growths gives insight as to what methods to choose.

Some surveys have characteristics which suggest the use of a particular method – for example, the necessity to stay within a tight schedule suggests CATI or CAWI because they are faster than PAPI. For the relevance of the analysis, the price series used is built entirely by varying: sample size, interview length and target incidence, without further hypothesis regarding budget or timing. Perfectly „typical" studies don't exist in reality, but we use this hypothesis to ensure price comparability and to help reveal the relation between variables (independent from other factors).

Market research agencies use calculation sheets to evaluate survey prices according to specifications. The final price includes the influence of statistical aspects (sample size etc.) as well as management aspects (indirect costs associated to the each data collection method).

We are trying to reveal in what measure the factors considered are influencing the final price. The method employed is Latin squares, a statistical method which deconstructs the influence of several factors on a variable, in order to find out which are the significant influence factors. The method was initially proposed by Fisher in "The Design of Experiments" and is based on variance analysis.

The prices are organized in a 3 × 3 Latin square, where the row indicates the questionnaire length, the column indicates the data collection method and the treatment is the target incidence. The Latin squares method leads to calculating the statistic F, which indicates a statistically significant influence if the value is greater than $F_{critic}(v_1, v_2)$, where $v_1$, $v_2$ = freedom degrees (values are found in tables). The initial data and calculations are presented in Annex 1.

The F value indicates ***significant price differences between methodologies*** *and that* ***the interview length significantly influences survey price***. Target incidence does not influence price, even though it can pose difficulties on project management; in other words, it may influence the project duration rather than the cost.

Next we analyzed the influence of each factor on the price, within each data collection method. Management costs should be comparable; therefore the method employed is again 3x3 Latin squares, varying the questionnaire length (row), sample size (column) and target incidence (treatment). The numeric results are presented in Annex 2.

### 3.1. PAPI

In PAPI research the most important element which affects the price is ***sample size.*** The previously mentioned aspects which make PAPI surveys more expensive than CATI seem to be multiplied as sample size grows.

On the other hand, the ***questionnaire length does not affect price significantly***. Long questionnaires (30-45 minutes) can be applied on paper without significantly increasing the research price. This is an important finding, especially compared to CATI research. Applying long questionnaires through the telephone is physically more tiring and can interrupt the respondents' daily routine; it is also easier

for the respondent to refuse on the phone or to simply hang up than it is to refuse an interviewer face to face (even after 20-25 minutes of the interview).

### 3.2. CATI

***The price of a CATI survey does not seem significantly affected by any of the factors*** taken into consideration. This result should be investigated in future research, even if only for checking purposes. It is possible that CATI research has particular characteristics, other than those presented here, which influence price and which must yet be discovered.

### 3.3. CAWI

The price of a CAWI survey is influenced by ***sample size and questionnaire length.*** Other studies show that questionnaire length also affects the quality of the collected information – if the interview is too long, the respondent manifests tiredness, lack of attention, answers given in haste or without reading the question and ultimately may give up the survey altogether (Merz, 2010).

Unlike telephone studies, surveys via the Internet entirely eliminate the influence of the interviewer. The respondent can therefore renounce the survey at any point, without having to explain the action to another person or to himself. Internet surveys face the double challenge of keeping the respondent's interest until the end of the questionnaire and to find enough respondents who complete the entire survey.

The main challenges of CAWI surveys in Romania are the Internet penetration rate, which is among the lowest in Europe, and the construction of a comprehensive, correct contact database. Developed countries like the United Kingdom can run nationally representative online surveys because they meet both conditions – and they benefit from the full set of CAWI advantages: savings at printing and mailing questionnaires, savings with interviewer and input operator wages, costs with the telephone line etc. These advantages enabled CAWI to reach its current worldwide popularity; this is also the direction in which Romanian market research is heading.

***Target incidence*** is taken into account when calculating price in practice, but results show that it ***does not significantly affect the pricing of any methodology***. Further research may focus on replacing it with another factor and recalculating statistical tests.

Table 3 summarizes the influence of various criteria on survey price using the three methodologies: PAPI, CATI, CAWI. Choosing a data collection method is often a compromise between objectives and costs (Voicu, 2008). This summary can be used as a practicual guide for choosing a methodology, depending on the importance of price and on survey characteristics.

*Table 3*

**Evaluating data collection methods after price analysis**

| Criterion \ Method | PAPI | CATI | CAWI |
|---|---|---|---|
| Questionnaire length | does not influence price | does not influence price | influences price |
| Sample size | influences price | does not influence price | influences price |
| Target incidence | does not influence price | does not influence price | does not influence price |

### 4. Which method is more suitable?

This question arises in every survey, but the only universally valid answer is: "it depends". In some cases, the data collection method is imposed by the client or by previous surveys, in which case the change of methodology can significantly affect key performance indicators.

If the choice of methodology is not restricted, one should keep in mind that for the current situation in Romania, ***CAWI research is most suitable for surveys with "special" targets***, which meet the profile of Internet users, in order to obtain relevant results. Such a survey is easy to recognize by the respondent profile. As a second option, these surveys can use CATI – Internet users are very likely to have a mobile or landline.

***CATI and PAPI surveys can be successfully conducted on "regular" targets as well, including nationally representative.*** Telephone interviewing is considered more advantageous than face to face interviewing. We will test this hypothesis by analyzing the two price series using the Mann-Whitney test.

The Mann-Whitney test is a non-parametric statistical test that evaluates whether two independent samples come from the same population. The test is run on ordinal variables. Here the two samples are the price series (presented in Annex 3). The test will show if there is a significant difference between prices (i.e. they come from different populations) or if the difference is not significant (i.e. they come from the same population).

The null hypothesis $H_0$ is: "prices are (statistically) equal"; the alternative hypothesis $H_1$: "there are significant differences between prices". The price values from the two series are brought together, are ordered from lowest to highest value and a rank is assigned to each ordered value (ranks from 1, 2 …).

The statistic of the test is $U_1 = R_1 - \frac{n_1(n_1+1)}{2}$, where $n_1$ is the number of observations in the first series and $R_1$ is the sum of ranks in the first series.

For each sample size: 500, 1000, 2000 respondents, the Mann-Whitney test shows ***there is a significant difference in price between CATI and PAPI surveys***.

Price ratio shows that a CATI survey is considerably less expensive than a PAPI one for short questionnaires (10 minutes) and at least medium target incidence (>10%). The cost of such a CATI survey can be up to 40-50% of PAPI cost. The greater the sample size, the more price-effective the use of CATI will be.

CATI is not recommended for long questionnaires (30-40 minutes) and medium or small incidence (<30%). Such a CATI survey can be up to 20-50% more expensive than a PAPI survey. The main factors explaining this situation are:

- Random Digit Dialing (RDD), which is used in CATI surveys to simulate the random selection of respondents and which offers statistically valid results. If the target incidence is low, the survey may require more time because there is no control over respondent selection;
- The lower desire of respondents to take part in telephone surveys in general (Henson et al., 1978) and to long telephone surveys in particular, aspect stated by the major researchers in this field (Groves, 1977; Groves and Kahn, 1979) and supported by recent studies (Holbrook et al., 2003).

As a final consideration, choosing a data collection method must not focus solely on price and overlook the research objectives. Each methodology has known risks which cannot be eliminated; therefore the choice must also consider how well the method can cover the objectives.

For example, a GfK NOP study on a nationally representative study of U.K. citizens aged 16 and older shows that the browser market is divided between Internet Explorer (59%), Firefox (29%) and Chrome (7%) (Patel, 2010). The study was conducted online and did not contain an explicit question regarding the browser used by the respondent; the program detected this setting automatically. The authors of the study present this method as offering more robust data than by respondents' explicit statement.

This study shows interesting, but rather unclear results, judging exclusively by the published specifications. Automatic browser detection shows the usage at a certain point in time, a "snapshot" of a momentary situation, "current usage" rather than "most often usage". It is possible that respondents answer the survey at work, at school, at a friend's house, at an Internet café etc., or even to intentionally use a different browser than the preferred one in order to ensure proper page display. The data series for Internet Explorer and Firefox shows variations up to ±2-3% compared to the previous month, which are difficult to explain on short term.

The research with automatic detection is ambiguous in terms of results. It collects correct information in the sense that it registers the correct browser, but it should be correlated and/or corrected with explicit usage or explicit preferences in order to reveal a wider, more relevant picture of the browser market.

## 5. Conclusions and recommendations

The previous analyses are based on real price calculations used by a market research company in 2010. The prices have been obtained by varying only the influence factors taken into consideration, all the other conditions remaining unchanged – to ensure data comparability and facilitate the search for significant influences.

### 5.1. Choosing a data collection method

Depending on the features and criteria of the research, some data gathering methods are more suitable than others. Considering the 4 characteristics included in the analysis, Table 4 summarizes the results of the research.

The results can be generalized for the entire field of market research because the cost evaluation for a research project is done in similar conditions for all companies within the field. Future research directions can cover: revealing the influence factors in CATI studies; including more factors in the price analysis; an analysis of costs actually recorded by the chosen research method (practice shows that costs are often different from the calculated price because of survey particularities).

*Table 4*

**Evaluating data collection methods depending on the features included in the analysis**

| Criterion \ Method | PAPI | CATI | CAWI |
|---|---|---|---|
| Price | Highest | Medium | Low |
| Representativity | National | National | Internet Users |
| Sample size | Large (>1000) | Large (>1000) | Small and medium (500-1000) |
| Questionnaire length | Including long questionnaires | No more than 30 minutes | No more than 25-30 minutes |
| Incidence of the target audience | Including low | Medium and high | Including low* |

** in Internet users.*

For evaluating the situation of other countries, we recommend the recalculation of the statistical tests based on data in each country. As noted before, Romania is an atypical European market; therefore conclusions should not be generalized at international level. The development level of each country influences the characteristics and behaviour of the consumer, as well as the level of technological equipment. This affects the types and prices of the studies used on each market.

Cultural differences in respondent reactions to data collection methods must also be taken into account, because they affect the quality of the results. A Bloomerce study shows differences in the way citizens from different countries respond to surveys: online research in Great Britain reflects the reality with great accuracy, while answers coming from Italy and Spain will reflect a somewhat distorted image. This happens because in online research, Italian and Spanish respondents were more likely to give socially desirable answers, extreme answers or to agree with certain statements (Forciniti and Groenland, 2003).

### 5.2. The future of data collection in Romanian market research

The results in this paper are applicable in Romania at present. In order to see how data collection methods will evolve in Romania, we can look towards Western Europe. Development is brought to Romania by international market research institutions, which bring here standards and advanced methodologies, developed in their home countries.

Most probably, during the up-coming years Romania will go through the same changes that the Western market research has already seen:

– *Telephonic surveys will gradually replace face to face research*, landline dialing giving way to mobile dialing (this tendency is also identified by Bălan, 2006). This evolution is possible and is in fact taking place during present times in Romania, because the existing infrastructure allows it. The penetration of fixed telephony in the population is 24.1%, while research shows that around 70% of Romanians have at least one SIM card (the penetration of mobile telephony active SIM cards is 113.5%). Improving the quality of telephone research compared to face to face research has to be evaluated when making this change.

– *The growing Internet penetration rate will enable the use of this communication mean to collect data for all types of studies* (Fine, 2000): quantitative ad-hoc, panel and qualitative surveys will be conducted online. In quantitative ad-hoc research the share of online surveys will grow and representativeness will increase to cover the entire population. The development of online surveys in 2010 in Romania mirrors that of telephone research in 2001. Bălan predicted at the time (correctly) that the share of CATI in Romanian research would grow.

– Other types of research will benefit from the Internet. Household panel research will become simpler and more accurate as paper diaries will be replaced by bar code scanning. Qualitative research will embrace online focus groups, which have advantages over classic focus groups (e.g. in international studies, there is no need to gather respondents physically in one location).

**Acknowledgement**

This article is a result of the project "Doctoral Program and PhD Students in the education research and innovation triangle". This project is co funded by European Social Fund through The Sectorial Operational Programme for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies.

**Notes**

[1] Kanetkar (2000) cites the results of a study which compared the quality of data obtained in 31 studies, where the response rate was 6% higher in the case of PAPI, but other variables did not register significant differences (missing values, the similarity of answer distributions etc.)
[2] Holbrook et al. (2003) argue that the use of CATI in long interviews leads to a higher non-response rate and to more answers "*indifferent*", more answers tending to agree - there is a desire to give socially acceptable answers.

## References

ANCOM (2009), *Raportul anual 2009 al Autorității Naționale pentru Administrare și Reglementare în Comunicații*, București

Bălan, C., "CATI: Tehnică de intervievare telefonică**"**, *Management & Marketing*, 1(1), 2006, pp. 81-92

Chan, P., "Sampling techniques", in Chakrapani, C. (editor), *Marketing research: state-of-the-art perspectives*, Chicago, 2000, pp. 156-184

Deblès, R., Teyssonnière de Gramont, E. (1991), *Études de marché*, 2nd Edition, Belfond, Paris

ESOMAR (2009), *Global Market Research 2009: ESOMAR Industry Report*, Amsterdam

Fine, B., "Internet Reasearch: The Brave New World", in Chakrapani, C. (editor), *Marketing research: state-of-the-art perspectives*, Chicago, 2000, pp. 143-155

Forciniti, N., Groenland, E., "Cross-cultural influences in European online marketing research", *Bloomerce White Paper*, No. 1, SSI/ Bloomerce, Rotterdam, 2003

Groves, R.M., "An Experimental Comparison of National Telephone and Personal Interview Surveys", *Proceedings of the Section on Social Statistics: American Statistical Association*, American Statistical Association, Washington, DC, 1977, pp. 232–241

Groves, R.M., Kahn, R.L. (1979), *Surveys by Telephone: A National Comparison with Personal Interviews*, Academic Press, New York

Henson, R., Cannell, C.F., Roth, A., "Effects of interview mode on reporting of moods, symptoms and need for social approval", *Journal of Social Psychology*, 105, 1978, pp. 123-129

Holbrook, A.L, Green, M.C., Krosnick, J.A., "Telephone versus face-to-face interviewing of national probability samples with long questionnaires", *Public Opinion Quarterly*, 67, 2003, pp. 79-125

Isaic-Maniu, A., Mitruț, C., Voineagu, V. (2004), *Statistică*, 2nd Edition, Editura Universitară, București

James, D. (GfK Marktforschung), "The future of online research", *Marketing News*, 1, 2000, p. 11

Kanetkar, V., "Data collection methods and marketing research: a comparison and review of alternatives", in Chakrapani, C. (editor), *Marketing research: state-of-the-art perspectives*, Chicago, 2000, pp. 106-142

Merz, J. "Online surveys in the YouTube Age", presented at *GfK Research Summit 2010*, June 14-15, 2010, Vienna

Patel, P., "Google continues the onslaught with Chrome", GfK TechTalk Blog, available at: http://www.gfktechtalk.com/2010/07/15/google-continues-the-onslaught-with-chrome/ (accessed July 21st, 2010)

Strong, C. (GfK NOP Technology), "Using mobile devices to get closer to consumers", presented at *GfK Research Summit 2010*, June 14-15, 2010, Vienna

Tull, D.S., Hawkins, D.I. (1990), *Marketing Research: Measurement & Method*, 5, Macmillan Publishing Company, New York

Voicu, M.C., "Utilizarea sondajului statistic în cercetarea preferințelor consumatorilor", *Revista Română de Statistică*, 7, 2008, p. 72

## Annex 1

This 3x3 Latin square contains prices in Euro for different combinations of variable values: data collection method (column), questionnaire length (row) and target incidence (treatment). This is not an exhaustive presentation of the possible cases; it is a summary of relevant combinations used in practice.

Calculation steps:

- Correction;
- sum of squares per row ($SP_R$), column ($SP_C$), treatment ($SP_{Tr}$) and total ($SP_T$), the sum of square errors ($SP_E$);
- mean squares per row ($MP_R$), column ($MP_C$), treatment ($MP_{Tr}$) and errors ($MP_E$).

The test statistic (the F value) is calculated for the row, column and treatment and represents the ratio between the mean squares for each element, divided by the mean of squares per errors. If an F value is greater than the value in tables, the element which registers this value has a significant influence on price.

High F values indicate that the row (questionnaire length) and column (data collection method) influences the price, while incidence does not affect it signficantly.

| Method / Q'nnaire length | CATI | PAPI | CAWI |
|---|---|---|---|
| 10 minutes | 6076 | 7687 | 4660 |
| 20 minutes | 8074 | 10337 | 6193 |
| 30 minutes | 9614 | 11353 | 8610 |

| incidence |
|---|
| <10% |
| 10-30% |
| 30-60% |

| | | | | | |
|---|---|---|---|---|---|
| correction | 586664914.7 | | | | |
| $SP_R$ | 20890648.72 | $MP_R$ | 10445324.36 | $F_{row}$ | 38.99 |
| $SP_C$ | 15509995.72 | $MP_C$ | 7754997.861 | $F_{column}$ | 28.95 |
| $SP_{Tr}$ | -572034.611 | $MP_{Tr}$ | -286017.306 | $F_{treatment}$ | -1.07 |
| $SP_T$ | 36900237 | $MP_E$ | 267906.6806 | | |
| $SP_E$ | 1071627 | | | | |

## Annex 2

Research prices are presented for each data collection method in part and analyzed using Latin squares. Particularities of each method are taken into consideration, for example CAWI studies are generally not as lengthy as PAPI and may use smaller samples (the Internet users are fewer than the general population, therefore smaller samples may be enough to ensure representativity).

### (a) PAPI

| Sample size / Q'nnaire length | 500 respon-dents | 1000 respon-dents | 1500 respon-dents | 2000 respon-dents |
|---|---|---|---|---|
| 10 minutes | 9016 | 12821 | 17777 | 25682 |
| 20 minutes | 10079 | 17394 | 19524 | 25719 |
| 30 minutes | 10448 | 18997 | 25679 | 28054 |
| 40 minutes | 11767 | 19560 | 27308 | 36287 |

| incidence |
|---|
| <10% |
| 10-30% |
| 30-60% |
| >60% |

| | | | |
|---|---|---|---|
| correction | 6245424784 | | |
| $SP_R$ | 124562726 | $MP_R$ | 41520908.67 |
| $SP_C$ | 750634614 | $MP_C$ | 250211538 |
| $SP_{Tr}$ | 17036230 | $MP_{Tr}$ | 5678743.333 |
| $SP_T$ | 922633172 | $MP_E$ | 5066600.333 |
| $SP_E$ | 30399602 | | |

| | |
|---|---|
| $F_{row}$ | 8.20 |
| $F_{column}$ | 49.38 |
| $F_{treatment}$ | 1.12 |

### (b) CATI

| Sample size / Q'nnaire length | 500 respon-dents | 1000 respon-dents | 2000 respon-dents | 3000 respon-dents |
|---|---|---|---|---|
| 10 minutes | 6076 | 8393 | 9265 | 16413 |
| 20 minutes | 8074 | 16136 | 14459 | 22560 |
| 30 minutes | 9614 | 19994 | 46296 | 30456 |
| 40 minutes | 10362 | 19746 | 47539 | 89932 |

| inci-dence |
|---|
| <10% |
| 10-30% |
| 30-60% |
| >60% |

| | | | |
|---|---|---|---|
| correction | 8803834327 | | |
| $SP_R$ | 2385151396 | $MP_R$ | 795050465.4 |
| $SP_C$ | 2323949433 | $MP_C$ | 774649811.1 |
| $SP_{Tr}$ | 511321076.7 | $MP_{Tr}$ | 170440358.9 |
| $SP_T$ | 7110402234 | $MP_E$ | 314996721.4 |
| $SP_E$ | 1889980328 | | |

| | |
|---|---|
| $F_{row}$ | 2.52 |
| $F_{column}$ | 2.46 |
| $F_{treatment}$ | 0.54 |

### (c) CAWI

| Sample size / Q'nnaire length | 200 respon-dents | 300 respon-dents | 500 respon-dents | 1000 respon-dents |
|---|---|---|---|---|
| 5 minutes | 2700 | 2640 | 3455 | 5780 |
| 10 minutes | 3120 | 3790 | 4110 | 6380 |
| 15 minutes | 3500 | 4239 | 5763.1111 | 8202 |
| 20 minutes | 4030 | 4872.5 | 6320 | 9920 |

| incidence |
|---|
| <10% |
| 10-30% |
| 30-60% |
| >60% |

| | | | |
|---|---|---|---|
| correction | 388298957.6 | | |
| $SP_R$ | 16297794.25 | $MP_R$ | 5432598.082 |
| $SP_C$ | 42397115.2 | $MP_C$ | 14132371.73 |
| $SP_{Tr}$ | 420371.4472 | $MP_{Tr}$ | 140123.8157 |
| $SP_T$ | 62030836.93 | $MP_E$ | 485926.0056 |
| $SP_E$ | 2915556.033 | | |

| | |
|---|---|
| $F_{row}$ | 11.18 |
| $F_{column}$ | 29.08 |
| $F_{treatment}$ | 0.29 |

## Annex 3

The Mann-Whitney test is based on ranks – the price series is ordered from lowest to highest value regardless of methodology, then each price is assigned a rank. The sum of ranks is calculated for each methodology and is tested against the reference value.

The test was calculated for three sample sizes: 500, 1000, 2000 respondents. With 95% probability we can accept that price differs significantly between CATI and PAPI studies even when survey specifications are the same.

| | Sample size | 500 respondents | | 1,000 respondents | | 2,000 respondents | |
|---|---|---|---|---|---|---|---|
| Questionnaire length | Incidence | PAPI | CATI | PAPI | CATI | PAPI | CATI |
| 10 minutes | <10 | 9016 | 6076 | 15652 | 9257 | 27776 | 15401 |
| | 10 - 30 | 8425 | 5001 | 14560 | 7273 | 25682 | 11843 |
| | 30 - 60 | 7687 | 4187 | 13155 | 6015 | 22940 | 9265 |
| | >60 | 7520 | 4002 | 12821 | 5493 | 22272 | 8393 |
| 20 minutes | <10 | 10337 | 9821 | 17394 | 16136 | 29885 | 28736 |
| | 10 - 30 | 10079 | 8074 | 16968 | 12619 | 29126 | 21719 |
| | 30 - 60 | 9174 | 6578 | 15228 | 9760 | 25719 | 16160 |
| | >60 | 8946 | 6232 | 14833 | 8959 | 24990 | 14459 |
| 30 minutes | <10 | 11611 | 14734 | 19421 | 25246 | 32949 | 46296 |
| | 10 - 30 | 11353 | 11956 | 18997 | 19994 | 32193 | 36074 |
| | 30 - 60 | 10448 | 9614 | 17257 | 15365 | 28783 | 26865 |
| | >60 | 10220 | 8303 | 16862 | 12736 | 28054 | 21596 |
| 40 minutes | <10 | 13157 | 18889 | 21722 | 33132 | 36287 | 61532 |
| | 10 - 30 | 12753 | 15395 | 20986 | 26139 | 34881 | 47539 |
| | 30 - 60 | 11995 | 12136 | 19560 | 19746 | 32119 | 34946 |
| | >60 | 11767 | 10362 | 19165 | 16050 | 31390 | 27500 |

| Sample size / Sum of ranks | 500 resp'ts | 1000 resp'ts | 2000 resp'ts |
|---|---|---|---|
| CATI | 235.5 | 229 | 231.5 |
| PAPI | 292.5 | 299 | 296.5 |

| | | | |
|---|---|---|---|
| U | 99.5 | 93 | 95.5 |
| $U_{0.05}$ | 75 | | |